Video Article

Ultrasound Images of the Tongue: A Tutorial for Assessment and Remediation of Speech Sound Errors

Jonathan L. Preston[1,2], Tara McAllister Byun[3], Suzanne E. Boyce[2,4], Sarah Hamilton[4], Mark Tiede[2], Emily Phillips[2], Ahmed Rivera-Campos[4], Douglas H. Whalen[2,5,6]

[1]Department of Communication Sciences and Disorders, Syracuse University

[2]Haskins Laboratories

[3]Department of Communicative Sciences and Disorders, New York University

[4]Department of Communication Sciences and Disorders, University of Cincinnati

[5]Program in Speech-Language-Hearing Sciences, City University of New York Graduate Center

[6]Department of Linguistics, Yale University

Correspondence to: Jonathan L. Preston at jopresto@syr.edu

URL: https://www.jove.com/video/55123
DOI: doi:10.3791/55123

Keywords: Behavior, Issue 119, ultrasound, speech therapy, tongue, articulation, biofeedback, rhotics, speech sound disorders

Date Published: 1/3/2017

Citation: Preston, J.L., McAllister Byun, T., Boyce, S.E., Hamilton, S., Tiede, M., Phillips, E., Rivera-Campos, A., Whalen, D.H. Ultrasound Images of the Tongue: A Tutorial for Assessment and Remediation of Speech Sound Errors. *J. Vis. Exp.* (119), e55123, doi:10.3791/55123 (2017).

Abstract

Diagnostic ultrasound imaging has been a common tool in medical practice for several decades. It provides a safe and effective method for imaging structures internal to the body. There has been a recent increase in the use of ultrasound technology to visualize the shape and movements of the tongue during speech, both in typical speakers and in clinical populations. Ultrasound imaging of speech has greatly expanded our understanding of how sounds articulated with the tongue (lingual sounds) are produced. Such information can be particularly valuable for speech-language pathologists. Among other advantages, ultrasound images can be used during speech therapy to provide (1) illustrative models of typical (*i.e.* "correct") tongue configurations for speech sounds, and (2) a source of insight into the articulatory nature of deviant productions. The images can also be used as an additional source of feedback for clinical populations learning to distinguish their better productions from their incorrect productions, en route to establishing more effective articulatory habits.

Ultrasound feedback is increasingly used by scientists and clinicians as both the expertise of the users increases and as the expense of the equipment declines. In this tutorial, procedures are presented for collecting ultrasound images of the tongue in a clinical context. We illustrate these procedures in an extended example featuring one common error sound, American English /r/. Images of correct and distorted /r/ are used to demonstrate (1) how to interpret ultrasound images, (2) how to assess tongue shape during production of speech sounds, (3), how to categorize tongue shape errors, and (4), how to provide visual feedback to elicit a more appropriate and functional tongue shape. We present a sample protocol for using real-time ultrasound images of the tongue for visual feedback to remediate speech sound errors. Additionally, example data are shown to illustrate outcomes with the procedure.

Video Link

The video component of this article can be found at https://www.jove.com/video/55123/

Introduction

Both clinical and research settings have seen an increase in the use of ultrasound imaging to provide visual biofeedback intervention to individuals with speech disorders. One important use of ultrasound imaging for speech-language pathologists is as a visual biofeedback tool during intervention for individuals with speech disorders. With the guidance of a speech-language pathologist, learners can observe real-time video of the shape and movements of their tongue and discuss how these images may differ from the tongue movements needed to properly articulate a speech sound. To conduct such interventions, it is important for users to be competent in the interpretation of ultrasound images as the tongue moves in real time. Knowledge of the range of correct articulatory patterns used by typical speakers is foundational to recognizing erroneous tongue shapes.

The methods described herein address (a) collecting ultrasound images of the tongue, (b) interpreting ultrasound images associated with both correct and incorrect productions of speech sounds, and (c) using real-time ultrasound imaging as a source of visual biofeedback to facilitate speech production changes in individuals with speech sound errors. Although ultrasound can be used to visualize a variety of lingual phonemes, examples here will focus on ultrasound images of the tongue for the /r/ sound (as in *red car*), which is described as the most common residual error among children acquiring American English [1]. It is also the sound that has been most extensively studied in clinical applications of ultrasound to date. [2-14]

One important goal in speech (re)habilitation is to facilitate more intelligible speech by teaching articulatory routines that result in perceptually appropriate productions of a target sound or sequence. Therefore, it is critical to understand tongue actions during normal speech and during production of speech errors. Real-time visualization of the tongue can play a highly beneficial role in encouraging a speaker to modify articulatory movements, as it provides the clinician and client with a shared representation of what is actually happening during speech. Without real-time visualization of the tongue, only static pictures or verbal descriptions of target tongue configurations are available to facilitate understanding of the desired articulatory behaviors. In schema-based models of motor learning, visual information about the movements of the tongue during speech is considered a form of "knowledge of performance" feedback (*i.e.* it provides specific qualitative information about the movement that occurred)[15]. Previous research has indicated that detailed knowledge of performance feedback can facilitate acquisition of a novel motor routine[16].

Ultrasound has several advantages over other technologies used to visualize speech. With ultrasound, the entire contour of the tongue can be visualized quickly from tip to root. Preparation for ultrasound imaging generally takes less than a minute.

In contrast, electropalatography (EPG) requires a dental impression and the creation of a customized pseudopalate (which may take weeks), and it can take time to adapt to speaking with the pseudo-palate [17]. EPG also enables visualization of tongue-palate contact only in the region covered by the pseudopalate and cannot display the tongue root or the overall shape of the tongue. This limits the nature of what aspects of articulation can be effectively targeted with EPG.

Another alternative is electromagnetic articulography (EMA), which can provide general information about tongue shape and movement [18]. However, EMA requires sensors to be glued to the tongue and other structures; thus, the set-up for this type of tongue imaging can take 20 - 30 min and may not be a viable method for frequent use. Thus, ultrasound may be viewed as more practical.

In the specific context of clinical research on the assessment and treatment of /r/ errors, the use of ultrasound has been reported in several studies for individuals with idiopathic speech sound disorders [2,10,11,13,19], hearing impairment [20], childhood apraxia of speech [12,21], and acquired apraxia of speech following a cerebral vascular accident [22]. Studies have also reported the use of ultrasound to treat errors on other lingual phonemes such as /s k g l ʃ ʧ / [23,24]. Additional populations that may be candidates include individuals with speech disorders related to cleft palate, or individuals learning pronunciation of sounds in a non-native language [25].

Ultrasound imaging may also be useful diagnostically, *e.g.*, to characterize errors in lingual shapes,[26,27], or to identify sub-perceptible or covert contrasts in disordered speech [28,29]. If precise articulatory measurements are being obtained and compared, it is essential that the ultrasound be stabilized so that the coordinate space for measurement remains reasonably constant. However, it is generally agreed that an unstabilized probe yields information of sufficient quality for clinical diagnosis and treatment applications, which is the focus of the present paper.

Protocol

Ethics Statement. When used in research, informed consent and/or assent from children is always required before collecting ultrasound images. When used clinically, clients should be informed of the purpose of the ultrasound imaging. Although diagnostic ultrasound imaging is considered "minimal risk" [30], users should always follow the ALARA principle when using ultrasound, meaning exposure to ultrasound should be as "As Low As Reasonably Achievable"[31]. This involves limiting acoustic power during imaging and also limiting exposure time. For example, if ultrasound is being used for visual feedback but the participant is not attending to the visual feedback, it would be prudent to discontinue imaging.

1. Collecting Ultrasound Images of the Tongue

NOTE: *Technical Considerations*. Diagnostic ultrasound probes are used to image the tongue. A frequency range between approximately 3 - 8 MHz with a frame rate of about 30 frames per second is recommended for clinical imaging the tongue [32].
NOTE: The instructions below apply to the diagnostic Ultrasound System (see **Materials Table**) with a C6-2 transducer, which was selected based on visual comparison of ultrasound images collected from several transducers available for this system. These instructions are adapted from the diagnostic ultrasound system reference manual for this device and are intended to be an illustrative example for one ultrasound. Many other ultrasound systems are in use, and users should consult the operating manuals of their specific device.

1. Turn on the machine. When powered on, observe the 2D imaging mode automatically displayed on the screen.
    1. Wait for the system to complete self-diagnostic and calibration tests.
    NOTE: The automatically displayed image can be adjusted to optimize the clinician's view by altering the settings of the instrument. Because the use of ultrasound for speech therapy is new, settings appropriate for imaging the tongue surface during speech are not typically preinstalled and must be installed by a representative from the manufacturer (preferred method), or by the clinician. It is important to have the right settings in order to safely and accurately image the tongue for speech therapy purposes.
    NOTE: Users should familiarize themselves with the basic operations of their ultrasound equipment be sure they understand how to make adjustments of all controls including depth, intensity and contrast to obtain the best images with their equipment. Power should be set as low as reasonably achievable for safety reasons, with adjustments in Gain to compensate for low power settings.
    2. To use system preset function, press the PRESETS key on the keyboard.
    3. Observe the Presets screen. Observe the menu items on the left and selections on the right of the screen.
    4. Roll the trackball to the menu item on the left of the screen and press the SET key on the control panel. Observe more selections for the selected menu item.
    5. To save changes and exit the system presets, press the SAVE button at the bottom of the screen.
    NOTE: **Table 1** shows the settings used for images in the video. Note that depth is transducer-dependent. Settings were developed in consultation with Siemens Corporation.
2. Place a small amount of ultrasound gel on the ultrasound transducer.
3. Position the participant comfortably in a chair with feet on the floor, back straight, and chin slightly forward.

4. To collect a sagittal image and visualize the tongue from tip to root, position the transducer vertically, making tight contact with the skin underneath the chin and applying a firm but not uncomfortable degree of pressure. Orient the transducer along the midline between the mental spine of the mandible and the hyoid bone.
5. Begin scanning with the ultrasound.
6. View the screen to verify that the transducer is properly oriented. In these sagittal images, the front of the tongue is on the right of the screen and the back of the tongue is on the left. Angle the ultrasound transducer slightly forward or slightly back depending on what part of the tongue is of interest.
NOTE: In some devices, the default settings will display an ultrasound image that is upside down. The user should check the default settings for their device and make adjustments as necessary.
7. Instruct the participant to swallow to orient the user to the tongue position relative to the palate.
8. Obtain a coronal view to image the tongue from left to right sides. To collect coronal images, rotate the ultrasound transducer 90 degrees. Instruct the participant to sustain sounds requiring midline grooving of the tongue such as /r, s, z, ʃ/.It may be necessary to adjust the transducer slightly further forward or back to visualize tongue grooving for different sounds.
9. Once imaging is complete, wipe off excess gel and clean the transducer with ultrasound-approved disinfecting wipes or ultrasound-approved cleansing spray.

2. Interpreting Ultrasound Images of the Tongue

1. **Basic interpretation of sagittal images**
    1. In a sagittal view, observe the tongue between two major shadows created by bone, which is opaque to ultrasound: the shadow of the mandible (anterior) and the shadow of the hyoid (posterior). At least one, and preferably both, of these landmarks are visible during tongue imaging.
    NOTE: If there is air below the tongue tip (as is typically the case when the tongue tip is extended), the extreme end of the tongue tip will not be visible on the ultrasound image.
    2. Instruct the participant to produce alveolar and velar sounds, such as /t d n/ then /k g/. This will help orient both the participant and the clinician to which side of the image is anterior/tongue tip and which side is posterior/tongue dorsum.

2. **Correct /r/ production**
    1. Instruct the participant to produce and sustain /r/. In a sagittal view of a correct production of /r/, the anterior portion of the tongue will elevate. The back part of the tongue dorsum slopes backwards for a correct /r/ production. Observe that if the probe is off midline or has rotated, the image will change and may be uninterpretable.
    NOTE: If the production is a "retroflex" /r/, then the tongue tip is raised toward the hard palate and the image of the tongue tip may be lost or may be represented as an artefact (*e.g.*, tongue tip appears to raise through the palate). In a classically "bunched" /r/, the tongue tip and/or blade are either horizontal or angled down toward the floor of the oral cavity, but the anterior portion of the tongue dorsum will be raised. In both cases, the elevated portion of the tongue is narrowing the airspace between the tongue and the palate-that is, it is making a vocal tract constriction. The location of this constriction gives English /r/ its primary palatal place of articulation.
    2. While the participant is sustaining /r/, visualize the tongue root.
    NOTE: Depending on the particular type of ultrasound probe being used, this may require angling the probe back toward the hyoid. Perceptually accurate English /r/ requires a secondary constriction in the vocal tract caused by the retraction of the tongue root toward the back pharyngeal wall. The posterior pharyngeal wall cannot typically be visualized with ultrasound, but retraction can be inferred: if the tongue root is retracted, the slope of the tongue surface behind the anterior constriction will be shallow.
    3. Rotate the probe 90 degrees to obtain a coronal view. Position the probe in the rough vicinity of the highest point of the tongue-about 1/3 of the distance between the chin and the throat in the sagittal plane.
    NOTE: While the participant is sustaining /r/, some elevation of the lateral margins of the tongue is generally visible. Although the teeth are not visible, it is common for individuals making a bunched type of /r/ to feel some contact with the back molars.

3. **Assessing distorted /r/ production with ultrasound**
    1. Instruct the participant to imitate and sustain /r/.
    NOTE: In a sagittal view of a distorted production of /r/, the anterior aspect of the tongue is typically low, the tongue dorsum usually is raised high and back. Tongue root retraction is often absent, which is often indicated by a steep or nearly vertical slope of the tongue shape behind the anterior constriction.
    2. Instruct the participant to imitate /r/ in a number of syllables, such as /ɝ, ɑr, ɪr, rɑ, ri, ru/. Note any contexts that are perceptually correct and identify tongue shapes that are associated with correct vs. incorrect productions.
    NOTE: Dialect differences will influence articulation; these examples are for American English.
    3. Rotate the probe 90 degrees to obtain a coronal view and observe the lateral margins of the tongue. Repeat productions of syllables such as /ɝ, ɑr, ɪr, rɑ, ri, ru/. While the participant is sustaining a distorted /r/, the lateral margins of the tongue often remain low on one or both sides.

3. Using Real-time Ultrasound Images for Feedback to Remediate Speech Sound Errors

1. Instruct the participant on proper positioning of the ultrasound. Allow the participant to hold the ultrasound transducer if they are able. Alternatively, have the clinical researcher hold the transducer, or clamp the transducer to a microphone stand to keep it steady while the participant leans forward and rests the chin on it, maintaining firm pressure.
NOTE: There may be a tendency for the probe to slide around. The clinician should be alert to correct movement of the probe to either side of the midsagittal plane, as the image becomes less consistent and harder to interpret.
2. Orient the participant to ultrasound images in sagittal section by teaching them about the parts of the tongue. The participant should be able to trace sample tongue contours on a sagittal ultrasound image. Instruct the participant that /r/ requires both an oral constriction in the front and a pharyngeal constriction in the back.

NOTE: It is often helpful to instruct the participant to identify the tip, blade, dorsum and root separately, as these correspond to independently moveable areas of the tongue.

3. Request that the participant point to the side of a real-time ultrasound image that represents the anterior and posterior, or "front of the tongue" and "back of the tongue."
4. Introduce the participant to different tongue shapes for /r/ using drawings, ultrasound images, or magnetic resonance images[33]. Make it clear where the oral and pharyngeal constrictions are, but also acknowledge that every tongue shape is slightly different.
NOTE: See 2.2.1 and 2.2.2 for interpreting images from correct /r/ production
5. Request that the participant describe 2 major constrictions for /r/ that are visible in sagittal section. If the participant cannot identify the oral and pharyngeal constrictions just described, continue to instruct.
NOTE: Participants should be able to report what is described in 3.4, indicating that some portion of the front of the tongue raises up, and that the tongue root moves back (as described in 2.2.1 and 2.2.2)
6. Orient the participant to ultrasound images in coronal section. Instruct the participant in the desired tongue shape from left to right. Require that the participant trace the tongue shape, identify the left and right edges, and center line groove. Request that the participant explain the desired shape.
7. Attempt to elicit correct /r/ in isolation or in syllables by providing phonetic cues to copy different tongue shapes.
NOTE: It can be helpful to provide "targets" on the screen for the participant to match during /r/ production. Targets can be generated using a cursor or by drawing on a transparency placed over the screen. Provide explicit instructions such as "If the tongue line doesn't match the red target, try doing something differently. Change the way one is saying the sound. Focus on making the dorsum go down," *etc.* See Supplement 1 for examples of Articulator Placement Cues.
8. Have the participant provide an explanation of what they did correctly or incorrectly after some of their attempts.
9. If necessary, use standard shaping strategies such as shaping /l/ to /r/ or /ɑ/ to /r/ while viewing the visual feedback.
10. After achieving correct productions, press the Pause button to freeze the image. Discuss how the correct image looks different from previous incorrect productions.
11. Practice throughout the session without the ultrasound to provide a break and to offer the opportunity for generalization to speech when visual feedback is not available.
NOTE: As the participant achieves more accurate productions, increase the linguistic complexity from syllables to words, phrases, and sentences, and reduce the amount of feedback to facilitate generalization.
NOTE: Accuracy of speech targets should be regularly monitored. This can be done perceptually by individuals (preferably who are blind to treatment status) who rate recordings of the sounds that are being trained with the ultrasound.

Representative Results

Figure 1 presents sample sagittal images of correct /r/ in a 9-year-old female. The ultrasound images are paired with magnetic resonance images from the same speaker to demonstrate the similar tongue shape that can be viewed with both technologies.

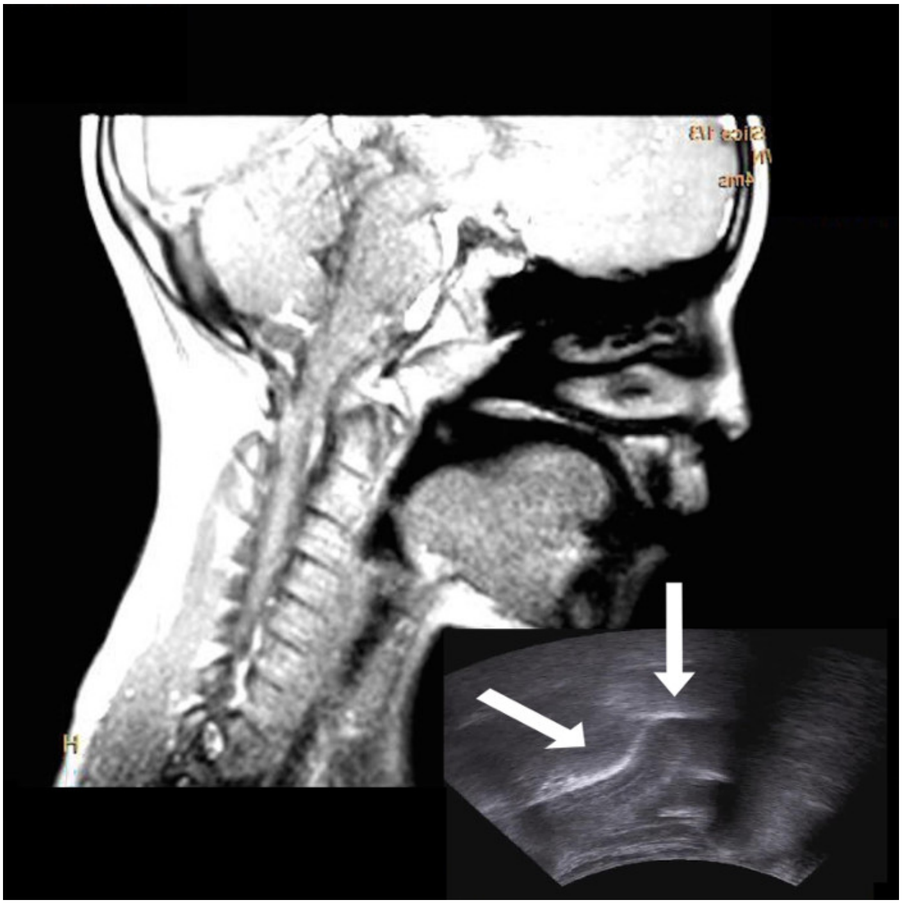

Figure 1: Sagittal View of a Magnetic Resonance Image during a Correctly Produced American English /r/ with Ultrasound Image of the Tongue (bottom right) from the Same Participant. In all images, the right side of the image represents anterior and the left represents posterior. Notice the elevation of the anterior tongue (right arrow) and the lowering of the dorsum (left arrow). Please click here to view a larger version of this figure.

In **Figure 2**, the same 9-year-old is shown 3 months earlier (before ultrasound visual feedback therapy). Note that the distorted /r/ involves a high posterior tongue position, low tongue tip/blade, and lack of a pharyngeal constriction, yielding a sound perceptually similar to [ʊ]. Correct /r/ productions feature elevation of the anterior tongue, a lowered tongue dorsum, and a posterior narrowing reflecting retraction of the tongue root. Note that a range of tongue shapes are possible for correct /r/.

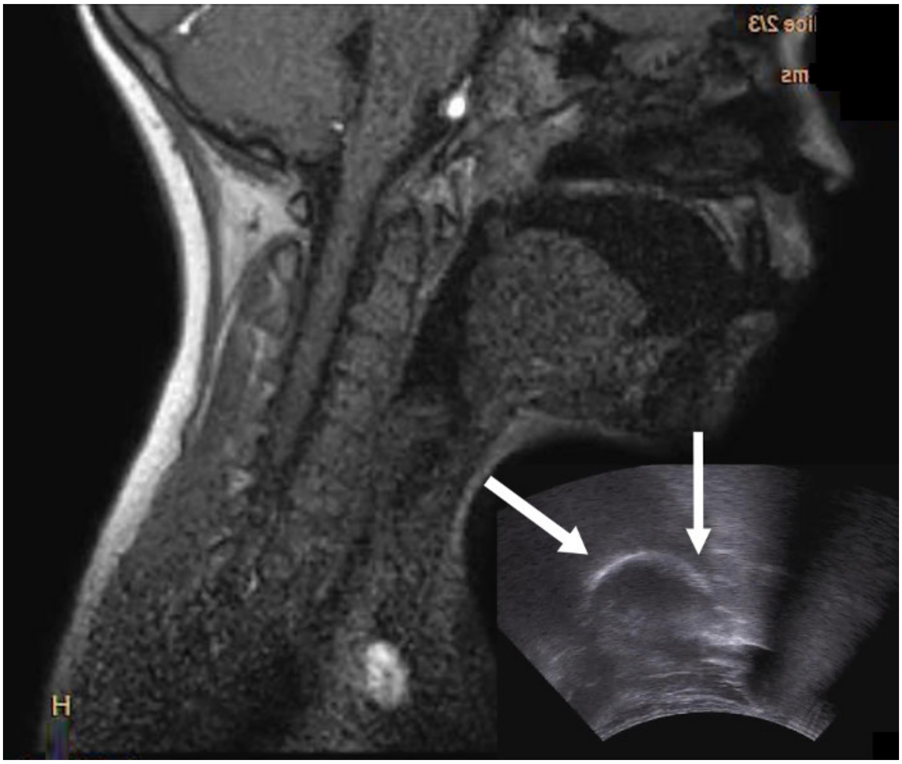

Figure 2: Sagittal View of a Magnetic Resonance Image during a Distorted Production of American English /r/ with Ultrasound Image of the Tongue (bottom right) from the Same Participant. In all images, the right side of the image represents anterior and the left represents posterior. Notice the low tongue tip/blade (right arrow) and the raised tongue dorsum (left arrow). Please click here to view a larger version of this figure.

Figure 3 shows sample correct and incorrect /r/ productions in coronal view. Note the elevation of the sides of the tongue, along with midline grooving, in the correct productions and a relatively flat tongue shape for distorted /r/.

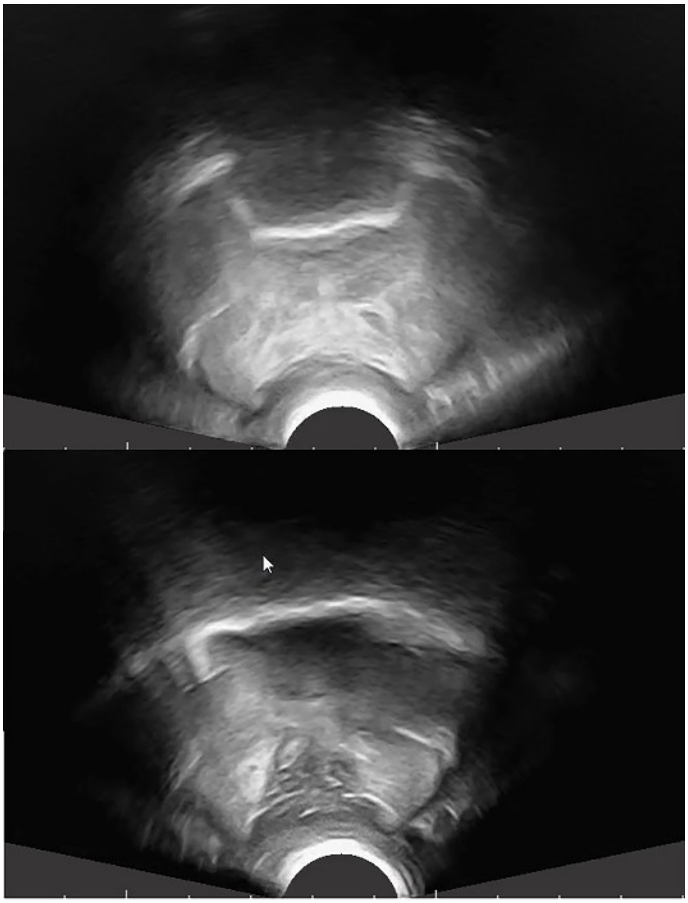

Figure 3: Sample Coronal Ultrasound Tongue Images of Correct (top) and Distorted (bottom) Productions of American English /r/ In these Coronal Views, the Probe is Positioned Vertically to Image the Posterior Tongue Dorsum. Notice the elevation of the lateral margins of the tongue for the correct /r/, along with a groove in the middle. Notice the flat tongue shape for the distorted /r/. These images are from an EchoBlaster 128 ultrasound. Please click here to view a larger version of this figure.

To date, studies on ultrasound visual feedback for speech sound errors have involved case series or single subject designs.[2,5,9-13,21-23] Widely varying patterns of individual response to treatment have been reported. For many individuals, improvement in sound accuracy can be observed with just a few hours of experimental treatment on /r/. Individuals who do not show immediate gains may still achieve improved production over the course of ultrasound practice. Gains made in the treatment setting almost always require some time to generalize to untreated words or contexts.

Figure 4 shows the average accuracy on words containing /r/ across 11 American English speaking participants ages 10-20 years who were treated for /r/ distortions. The data are from multiple-baseline across-subjects single case designs [13,34]. Some of the participants were treated on other sounds as well, although the figure is restricted to accuracy of /r/ in one word position per participant. The vertical axis represents percent of untreated /r/ words judged as correct. The horizontal axis represents separate sessions (spaced approximately 3 - 4 d apart) in which data were collected. Accuracy of /r/ production at the word level was monitored before, during, and after the 7 treatment sessions. Multiple listeners rated recorded productions of words as either "correct /r/" or "incorrect /r/" based on perceived phonetic accuracy. The box reflects the 7 sessions in which ultrasound biofeedback therapy was provided. Improved /r/ accuracy corresponds with the onset of treatment. Moreover, after 7 sessions, when treatment was withdrawn, an upward trend in accuracy continues, suggesting that retention and generalization continued to occur.

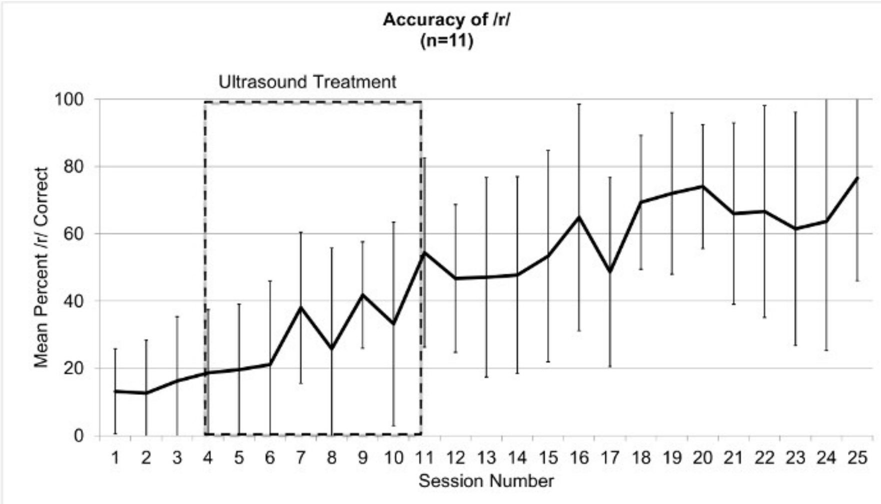

Figure 4: Mean Accuracy of /r/ in Single Words for 11 Participants Ages 10 - 20 Years Treated for /r/ Distortions. The box represents the sessions in which ultrasound visual feedback treatment occurred. Error bars represent standard deviations. Please click here to view a larger version of this figure.

Discussion

Critical Steps within the Protocol

It is essential to obtain clear, interpretable images as described in steps 1.3 and 1.6. Poor image quality renders the procedures meaningless. Additionally, participants must be fully aware of what they are seeing on the screen. Therefore, orienting the participant to the image as described in 3.2 is a step that should be emphasized prior to providing visual feedback training. Additionally, step 3.10, which involves clearly describing differences in tongue shape between the participant's perceptually accurate and inaccurate tongue shapes, is a critical step to increase awareness of the target tongue shape for a specific speaker.

Modifications and Troubleshooting

Image quality is essential. When image quality is waning, it may be necessary to re-apply gel and/or to check that the probe is making stable contact with the skin.

Additionally, it is important to recognize when the images are not representing what the user intends. For example, when collecting sagittal images, if the probe is positioned in the midsagittal plane (*i.e.,* down the middle of the head), the image will show the groove running down the center line of the tongue. If the probe is positioned to the side, the image will show more of the lateral edge of the tongue. The gross shape of the ultrasound "bright white line" will be similar if the image shows more of the groove or more of the tongue side, but they will not be exactly the same. The user should therefore regularly check the position of the probe to determine whether the images reflect mid-sagittal images, and reposition the probe if necessary.

Limitations of the Technique

Although ultrasound has significant advantages over other approaches to visualizing speech production, it is not without limitations. One primary limitation of ultrasound imaging is that only the tongue is imaged. That is, other structures such the hard or soft palate or the pharyngeal walls are not visible; thus, the relation of the tongue to other structures is not apparent. Additionally, it can be difficult to determine where exactly along the tongue contour the images are collected. For example, when interpreting sagittal images of the tongue, the position of the probe is important to consider, as images may not necessarily be mid-sagittal (*i.e.,* midline) if the probe is offcenter or has been rotated. Additionally, not all participants/clients tolerate the use of ultrasound gel beneath the chin. The mindful user of ultrasound should be aware of both the advantages and the limitations of the technology.

Significance of the Technique with Respect to Existing/Alternative Methods

Ultrasound imaging of the tongue using diagnostic mode can be a fast, safe, and effective technology for visualizing tongue movements in real time [30,32]. This information can be used to contrast correct and incorrect productions of speech sounds as a way to understand speech errors and teach desired movements for a variety of speech sounds. Traditional speech therapy methods for assessing and remediating speech sound errors such as /r/ distortions rely on auditory perception. Thus, the speech-language clinician is unaware of the exact nature of the speaker's tongue movements. Cues are often provided instructing speakers to modify their tongue position without any visual reference to the actual movement. Thus, real-time imaging of the tongue in offers an immediate visualization for shared discussion of speech, which traditionally has been abstract or transient. With respect to current theories on speech motor learning (*e.g.*, schema-based motor learning), ultrasound visual feedback offers a form of knowledge of performance feedback [13,15]. This feedback may facilitate the acquisition of new speech motor plans for individuals who have previously had difficulty understanding the target movements.

Ultrasound imaging can be particularly useful for evaluating [26,27] and remediating [2,10,11,12,13,20] speech sound errors that involve the oral and pharyngeal constrictions associated with /r/. Sagittal views can identify if the participant is lacking an anterior constriction or tongue root retraction. Coronal views provide the ability to examine whether there is midline grooving and elevation of the lateral margins of the tongue during /r/ production. Once the elements in error have been properly identified, this information can be used to systematically train new tongue movements, ideally while viewing real-time feedback of the tongue [2,10,11,12,13,20]. Methods such as electropalatography or electromagnetic articulography do not allow sufficient visualization of all aspects of the tongue, such as the tongue root, whereas ultrasound can overcome this limitation.

Future Applications or Directions after Mastering This Technique

The protocol outlined here is intended to be broad enough to allow others to follow the procedures regardless of the ultrasound technology available. The procedures are also intended to be flexible enough to meet a variety of clinical research or clinical practice needs. Although the focus throughout this discussion was on the specific context of treatment for /r/, these procedures can readily be adapted when training other speech sounds or when working with a variety of populations. Ultrasound feedback of the tongue can be useful for remediation of lingual sounds other than /r/, including vowels, velar and alveolar stops and nasals, and lingual fricatives and affricates [21,23].

Variations in procedures exist; for example, some researchers have used head stabilization techniques to prevent movement of the vocal tract relative to the ultrasound probe. Such procedures are useful if one intends to measure the contour of the tongue [23,35,36] and stabilization can also overcome some of the problems such as the drift in the position of the probe over time; however, head stabilization during ultrasound imaging of the tongue can lead to practical limitations (*e.g.*, uncomfortable head-mounted devices) and thus the ultrasound user must make decisions about the relative trade-off of such procedures. Studies are underway exploring specific modifications to the procedures (*e.g.*, the amount of practice with ultrasound that is ideal, the role of cueing only oral constrictions vs. oral and pharyngeal constrictions) to determine the methods that are optimally effective. In sum, evidence continues to accumulate that procedures incorporating ultrasound feedback of the tongue can yield improved speech clarity in individuals with speech sound disorders.

Disclosures

Siemens Corporation provided a temporary loan of three Acuson X300 ultrasounds for research purposes at no cost to the authors.

Acknowledgements

The work was supported by NIH grants R01DC013668 (D. Whalen, PI) and R03DC013152 (J. Preston, PI).

References

1. Ruscello, D. M. Visual feedback in treatment of residual phonological disorders. *J Commun Disord.* **28** 279-302 (1995).
2. Adler-Bock, M., Bernhardt, B., Gick, B., & Bacsfalvi, P. The Use of Ultrasound in Remediation of North American English /r/ in 2 Adolescents. *Am J Speech Lang Pathol.* **16** (2), 128-139 (2007).
3. Bacsfalvi, P., & Bernhardt, B. M. Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: ultrasound and electropalatography. *Clin Linguist Phon.* **25** (11-12), 1034-1043 (2011).
4. Bacsfalvi, P., Bernhardt, B. M., & Gick, B. Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. *Int. J. Speech Lang. Pathol.* **9** (1), 36-45 (2007).
5. Bernhardt, B. *et al.* Ultrasound as visual feedback in speech habilitation: Exploring consultative use in rural British Columbia, Canada. *Clin Linguist Phon.* **22** (2), 149 - 162 (2008).
6. Bernhardt, B., Bacsfalvi, P., Gick, B., Radanov, B., & Williams, R. Exploring the Use of electropalatography and ultrasound in speech habilitation. *Can. J. Speech Lang. Pathol.* **29** (4), 169-182 (2005).
7. Bernhardt, B., Gick, B., Bacsfalvi, P., & Adler-Bock, M. Ultrasound in speech therapy with adolescents and adults. *Clin Linguist Phon.* **19** (6/7), 605-617 (2005).
8. Bernhardt, B., Gick, B., Bacsfalvi, P., & Ashdown, J. Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clin Linguist Phon.* **17** (3), 199 - 216 (2003).
9. Fawcett, S., Bacsfalvi, P., & Bernhardt, B. Ultrasound as visual feedback in speech therapy for/r/with adults with Down syndrome. *Down Syndrome Quarterly.* **10** (1), 4-12, (2008).
10. Modha, G., Bernhardt, B., Church, R., & Bacsfalvi, P. Case study to use ultrasound to treat /r/. *Int J Lang Commun Disord.* **43** (3), 323-329 (2008).
11. McAllister Byun, T., Hitchcock, E. R., & Swartz, M. T. Retroflex versus bunched in treatment for rhotic misarticulation: Evidence from ultrasound biofeedback intervention. *J Speech Lang Hear Res.* **57** (6), 2116-2130 (2014).
12. Preston, J. L., Maas, E., Whittle, J., Leece, M. C., & McCabe, P. Limited acquisition and generalisation of rhotics with ultrasound visual feedback in childhood apraxia. *Clin Linguist Phon.* **30** (3-5), 363-381 (2016).
13. Preston, J. L. *et al.* Ultrasound visual feedback treatment and practice variability for residual speech sound errors. *J Speech Lang Hear Res.* **57** (6), 2102-2115 (2014).
14. Sjolie, G. *Effects of Ultrasound as Visual Feedback of the Tongue on Generalization, Retention, and Acquisition in Speech Therapy for Rhotics.* Masters thesis, Syracuse University, (2015).
15. Maas, E. *et al.* Principles of motor learning in treatment of motor speech disorders. *Am J Speech Lang Pathol.* **17** (3) (2008).
16. Newell, K., Carlton, M., & Antoniou, A. The interaction of criterion and feedback information in learning a drawing task. *J Mot Behav.* **22** (4), 536-552 (1990).
17. McLeod, S., & Searl, J. Adaptation to an electropalatograph palate: Acoustic, impressionistic, and perceptual data. *Am J Speech Lang Pathol.* **15** (2), 192-206 (2006).
18. Katz, W. *et al.* Opti-speech: A real-time, 3D visual feedback system for speech training. in *Proc. Interspeech.* (2014).

19. Shawker, T. H., & Sonies, B. C. Ultrasound Biofeedback for Speech Training: Instrumentation and Preliminary Results. *Invest Radiol.* **20** (1), 90-93 (1985).
20. Bacsfalvi, P. Attaining the lingual components of /r/ with ultrasound for three adolescents with cochlear implants. *Can. J. Speech Lang. Pathol.* **34** (3), 206-217 (2010).
21. Preston, J. L., Brick, N., & Landi, N. Ultrasound biofeedback treatment for persisting childhood apraxia of speech. *Am J Speech Lang Pathol.* **22** (4), 627-643, (2013).
22. Preston, J. L., & Leaman, M. Ultrasound visual feedback for acquired apraxia of speech: A case report. *Aphasiology.* **28** (3), 278-295 (2014).
23. Cleland, J., Scobbie, J. M., & Wrench, A. A. Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clin Linguist Phon.* **29** (8-10), 575-597 (2015).
24. Lipetz, H. M., & Bernhardt, B. M. A multi-modal approach to intervention for one adolescent's frontal lisp. *Clin Linguist Phon.* **27** (1), 1-17 (2013).
25. Gick, B. *et al.* Ultrasound imaging applications in second language acquisition. *Phonology and second language acquisition.* **36** 315-328 (2008).
26. Gick, B. *et al.* A motor differentiation model for liquid substitutions in children's speech. *Proceedings of Meetings on Acoustics.* **1** (1) (2007).
27. Klein, H. B., McAllister Byun, T., Davidson, L., & Grigos, M. I. A Multidimensional Investigation of Children's /r/ Productions: Perceptual, Ultrasound, and Acoustic Measures. *Am J Speech Lang Pathol.* **22** (3), 540-553 (2013).
28. Zharkova, N., Gibbon, F. E., & Lee, A. Using ultrasound tongue imaging to identify covert contrasts in children's speech. *Clin Linguist Phon.* 1-14 (2016).
29. McAllister Byun, T., Buchwald, A., & Mizoguchi, A. Covert contrast in velar fronting: An acoustic and ultrasound study. *Clin Linguist Phon.* **30** (3-5), 249-276 (2016).
30. Epstein, M. A. Ultrasound and the IRB. *Clin Linguist Phon.* **19** (6-7), 567-572 (2005).
31. Barnett, S. B. *et al.* International recommendations and guidelines for the safe use of diagnostic ultrasound in medicine. *Ultrasound Med Biol.* **26** (3), 355-366 (2000).
32. Lee, S. A. S., Wrench, A., & Sancibrian, S. How To Get Started With Ultrasound Technology for Treatment of Speech Sound Disorders. *SIG 5 Perspectives on Speech Science and Orofacial Disorders.* **25** (2), 66-80 (2015).
33. Boyce, S. E. The articulatory phonetics of /r/ for residual speech errors. *Seminars in Speech and Language.* **36** (4), 257-270 (2015).
34. Preston, J. L., Leece, M. C., & Maas, E. Motor-based treatment with and without ultrasound feedback for residual speech sound errors. *International Journal of Language & Communication Disorders.* (in press). (2016).
35. Cleland, J., Mccron, C., & Scobbie, J. M. Tongue reading: Comparing the interpretation of visual information from inside the mouth, from electropalatographic and ultrasound displays of speech sounds. *Clin Linguist Phon.* **27** (4), 299-311 (2013).
36. Zharkova, N., Gibbon, F. E., & Hardcastle, W. J. Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation. *Clin Linguist Phon.* **29** (4), 249-265 (2015).

